

← Speaker: Seong Joon Oh (NAVER)

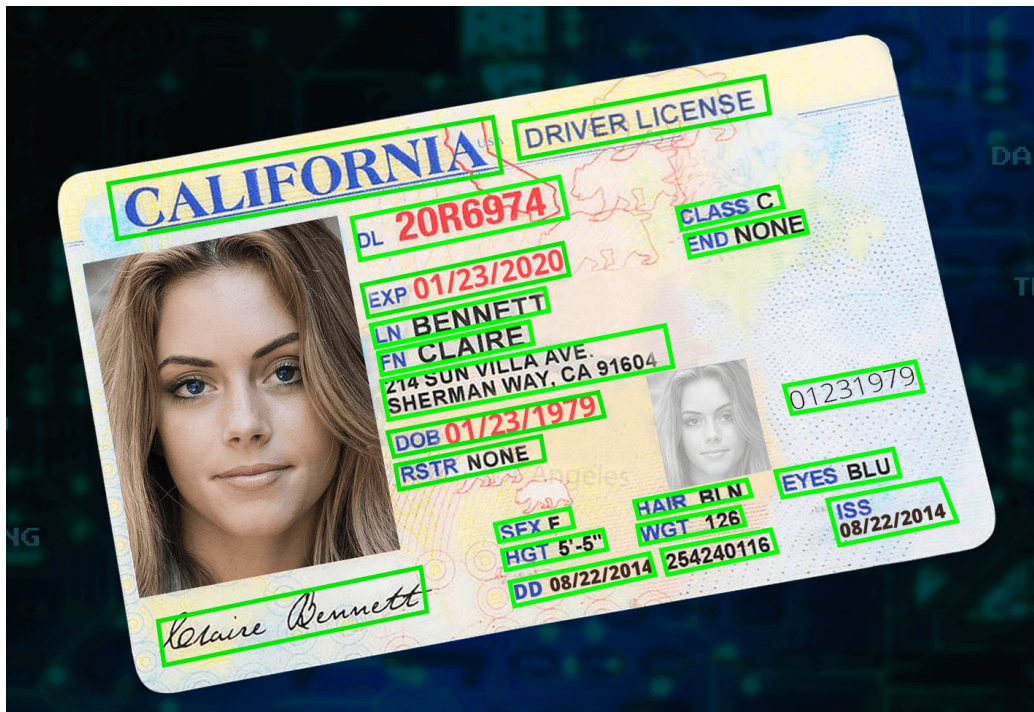
# Weakly-supervised learning: Conclusion

# Computer vision is finally working!



<https://cloud.google.com/vision/automl/object-detection/docs>

# Computer vision is finally working!

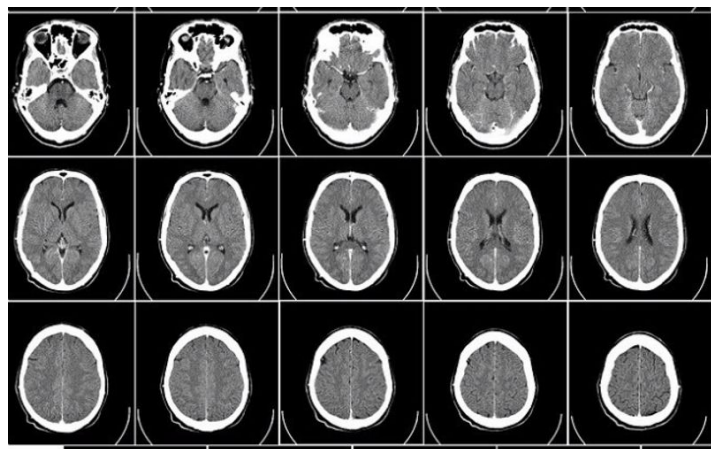


# Computer vision is finally working!



<https://www.biometricupdate.com/202001/airport-biometrics-predictions-deployments-upgrades-and-plans-for-future-services>

# Computer vision is finally working!



Medical AI



Self-driving cars

# Behind the success... Huge annotation costs.



**ImageNet1K**

Multi-label  
annotation takes

26 seconds / image



# Behind the success... Huge annotation costs.



**Cityscapes**

1.5 hour / image.

Cordts et al. The Cityscapes Dataset for Semantic Urban Scene Understanding. CVPR 2016.

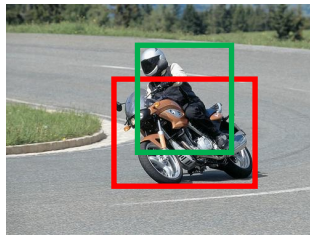
# Weak supervision: How to save costs on labelling.



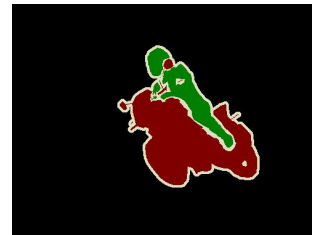
{motorbike, person}



{motorbike (point),  
person (point)}



{motorbike (b-box),  
person (b-box)}



{motorbike (pixel labels),  
person (pixel labels)}

1 sec  
per class

2.4 sec  
per instance

10 sec  
per instance

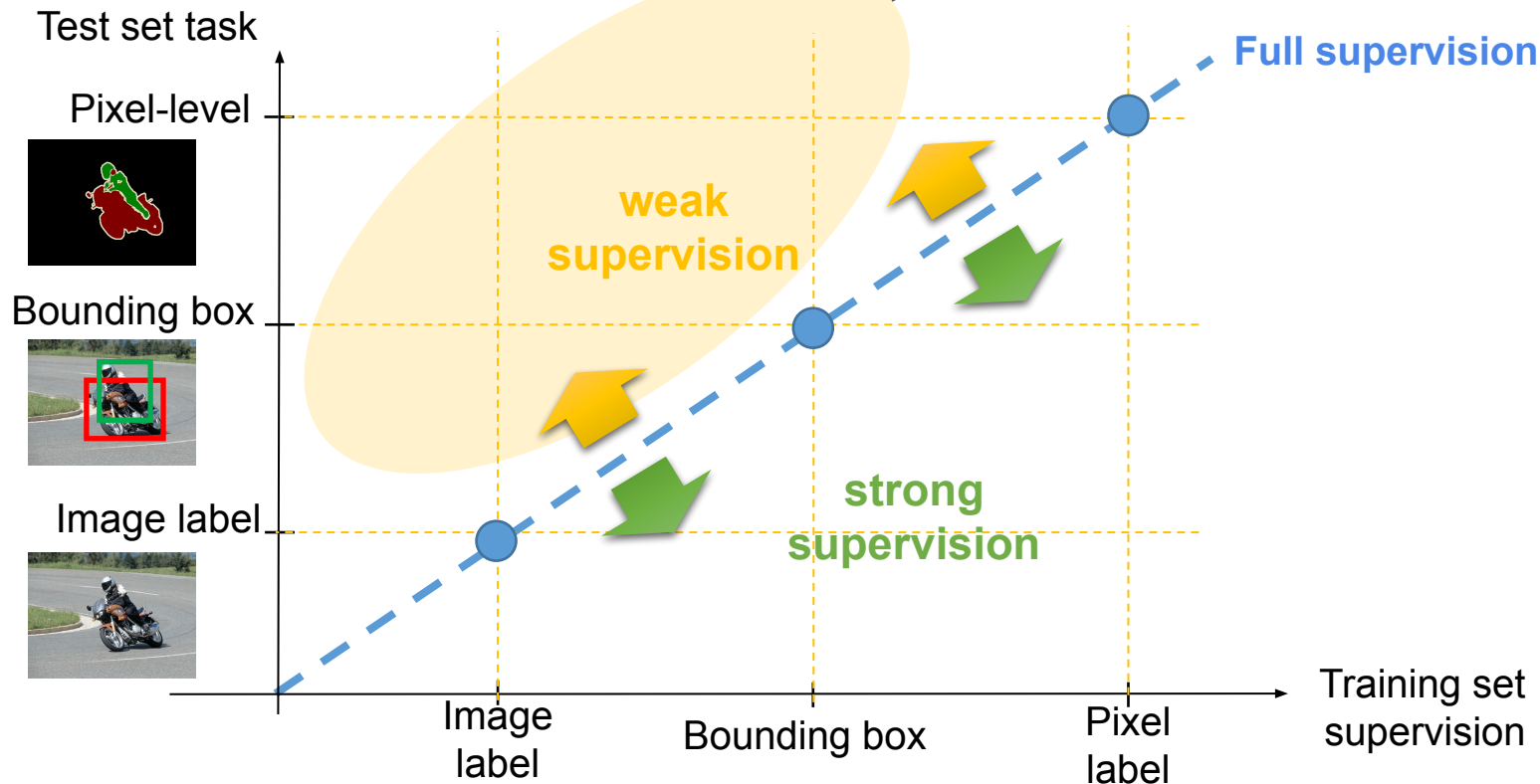
78 sec  
per instance

annotation time

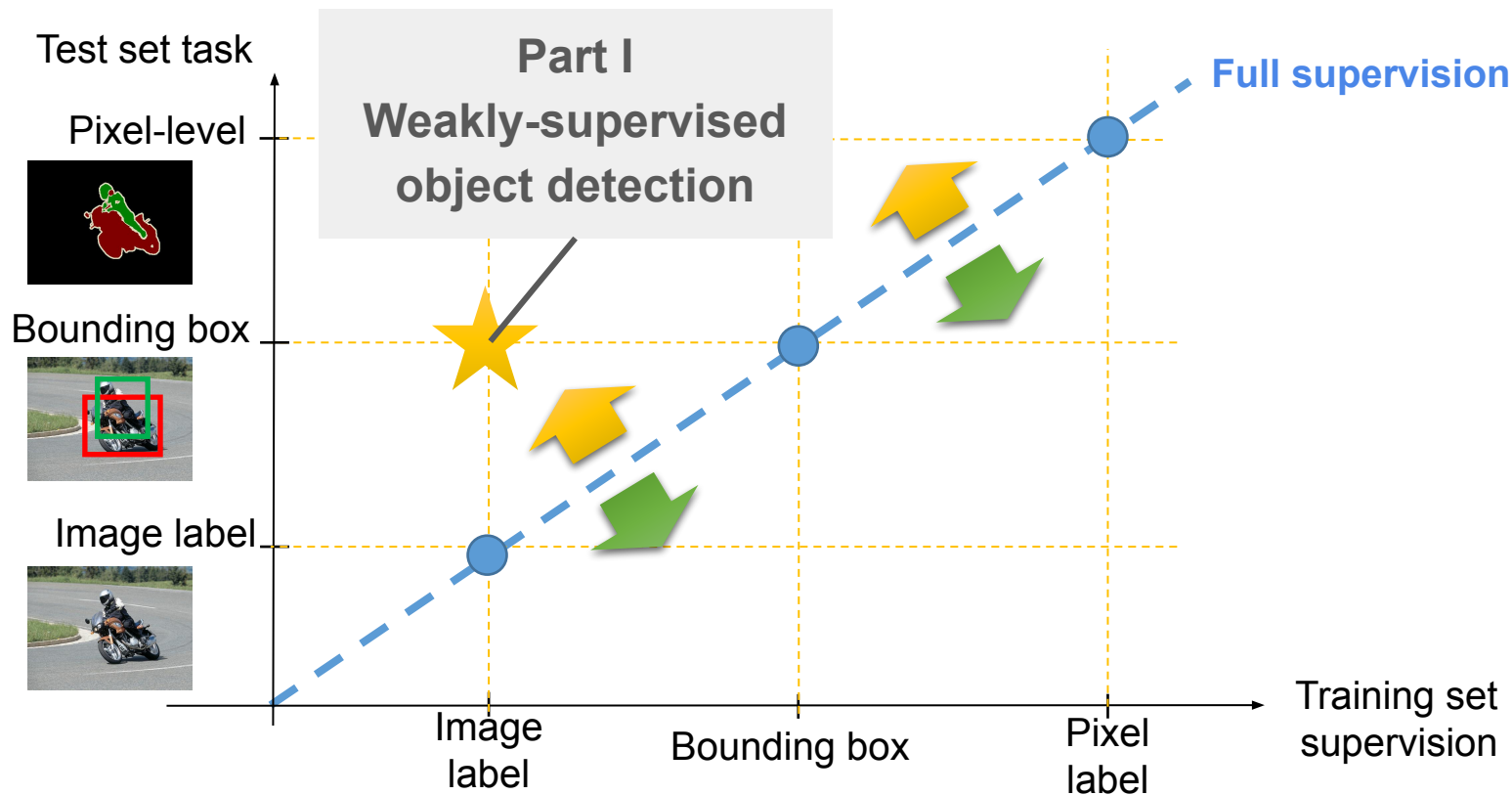


# Schematics of supervision types.

Covered by this  
tutorial

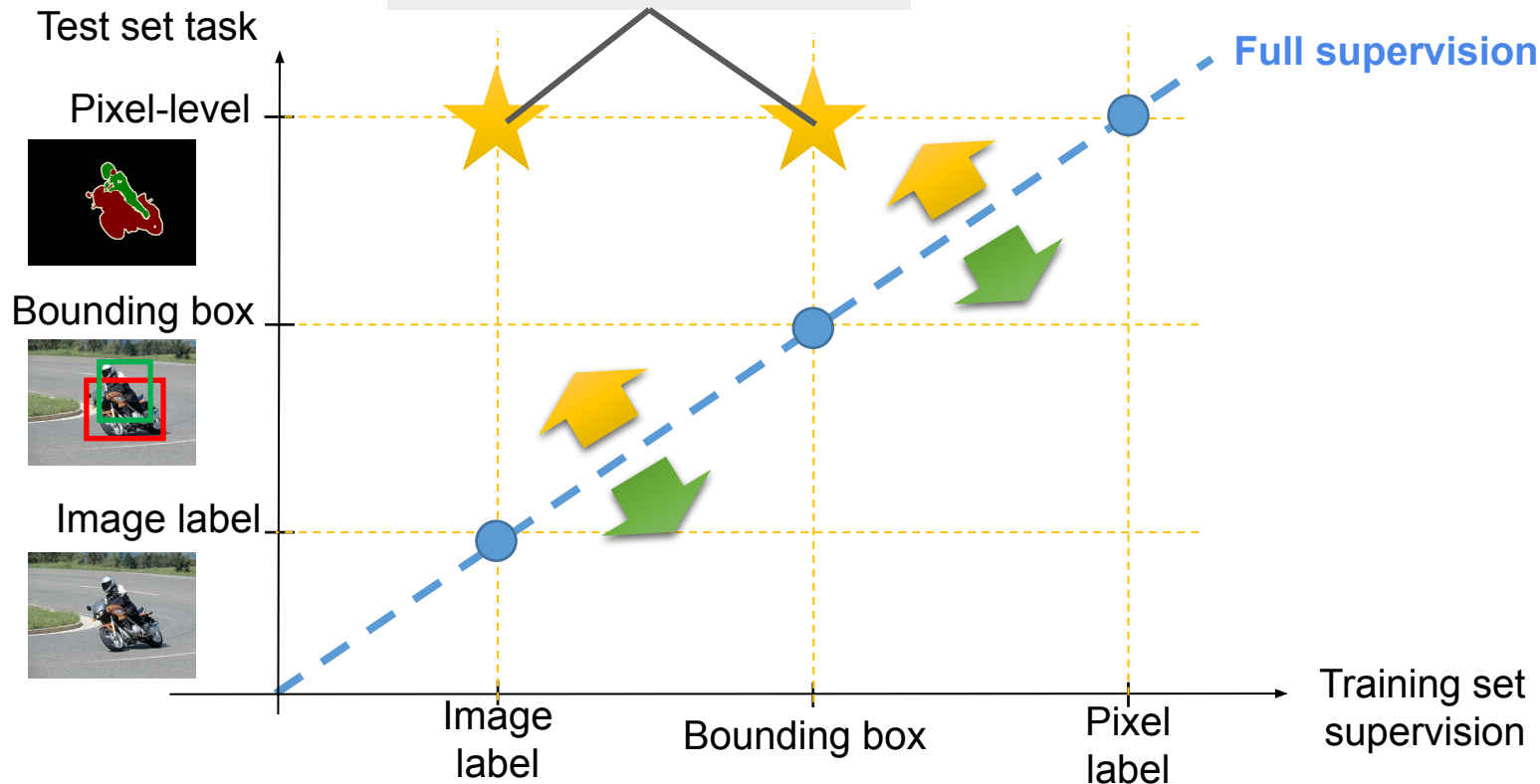


# Recap of tutorial.



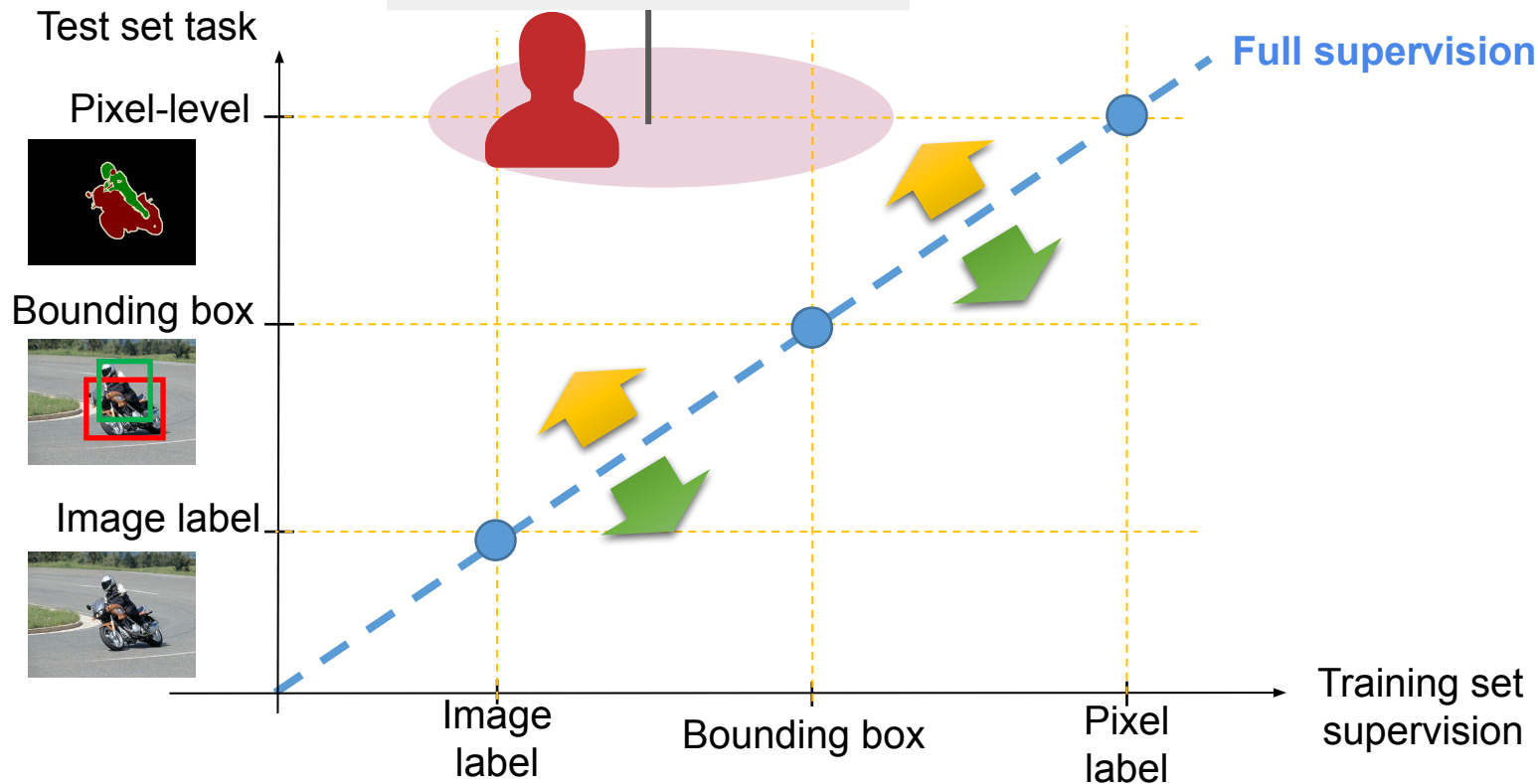
# Recap of tutorial

## Part II Weakly-supervised pixel labelling



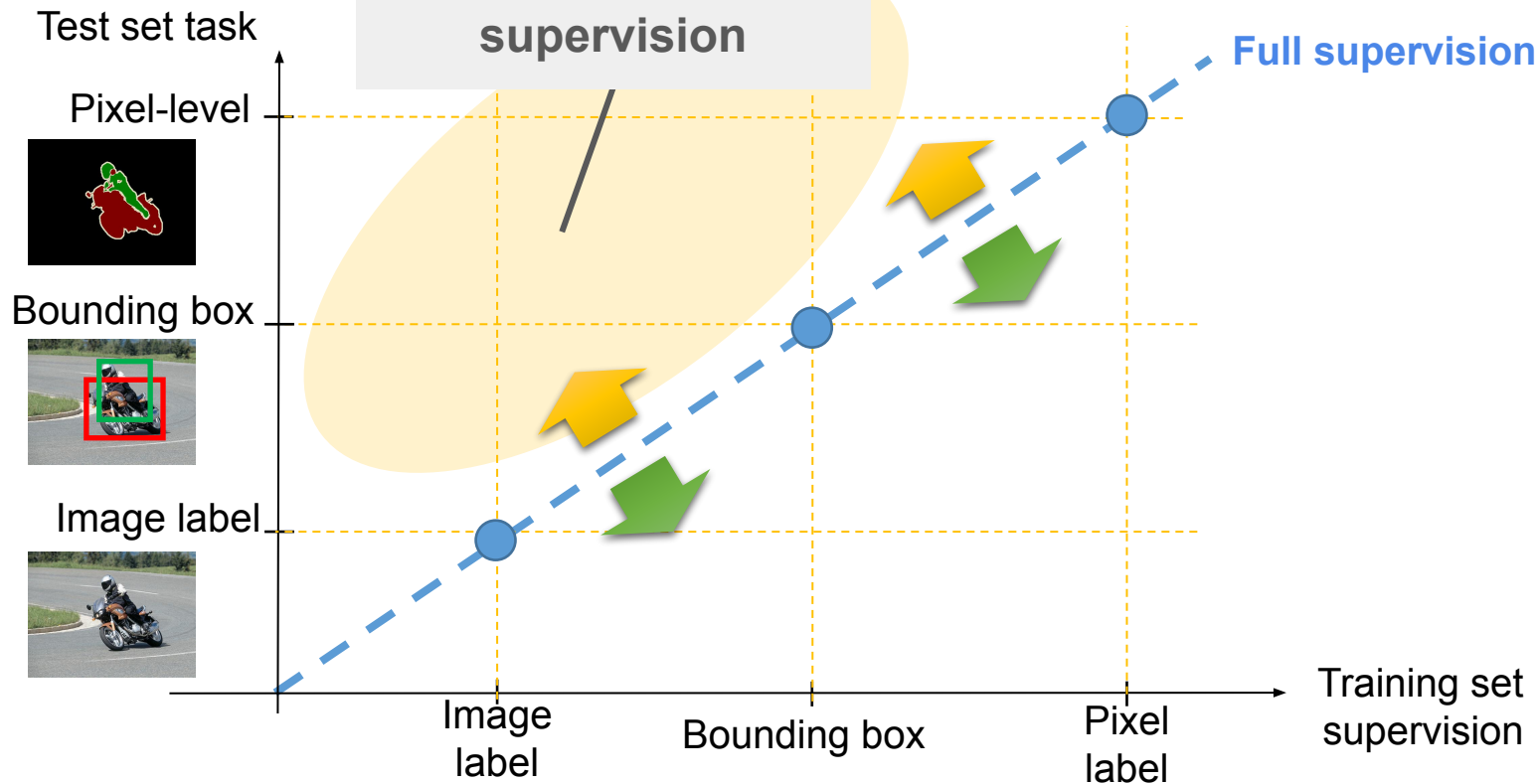
# Recap of tutorial

## Part III Human-in-the-loop annotation

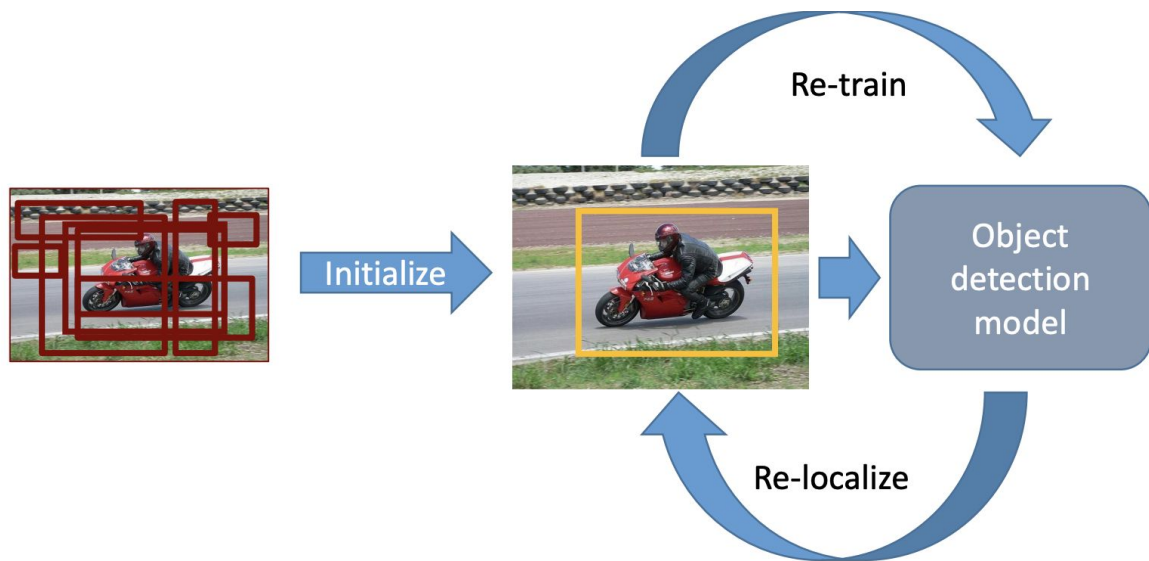


# Recap of tutorial

## Part IV Evaluation of weak supervision



# Overarching challenge: ill-posedness.



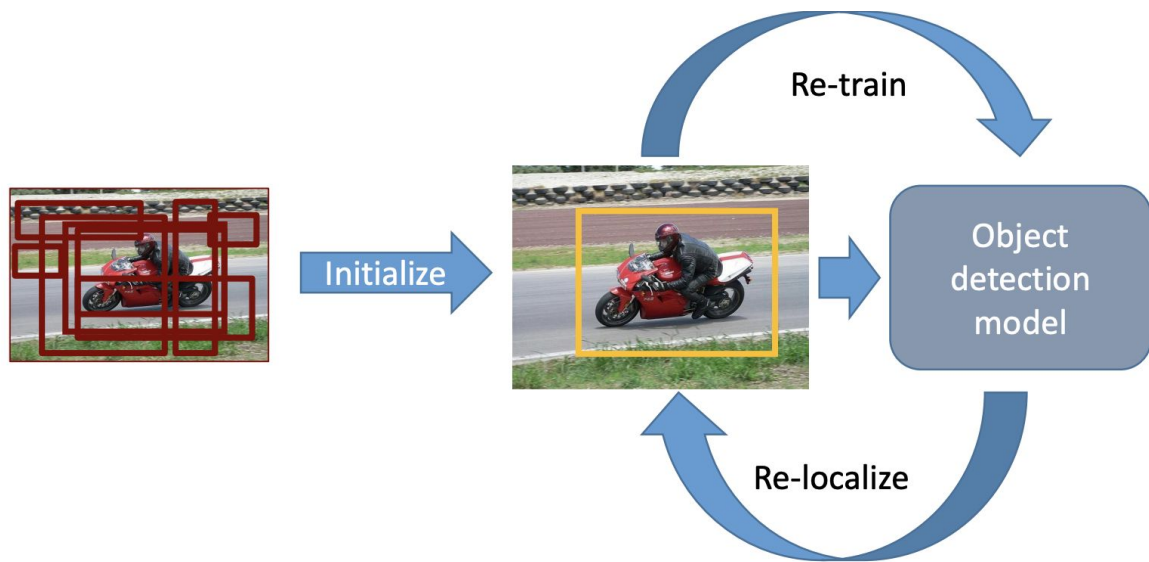
## Weakly-supervised object detection:

- Sensitivity to initialization.
- Many local minima.

→ Need priors & extra sources of information.



# Overarching challenge: ill-posedness.



**WSOD** + Prior information:

- Objectness
- Context
- Boundaries
- Scale
- Motion
- Equivariance

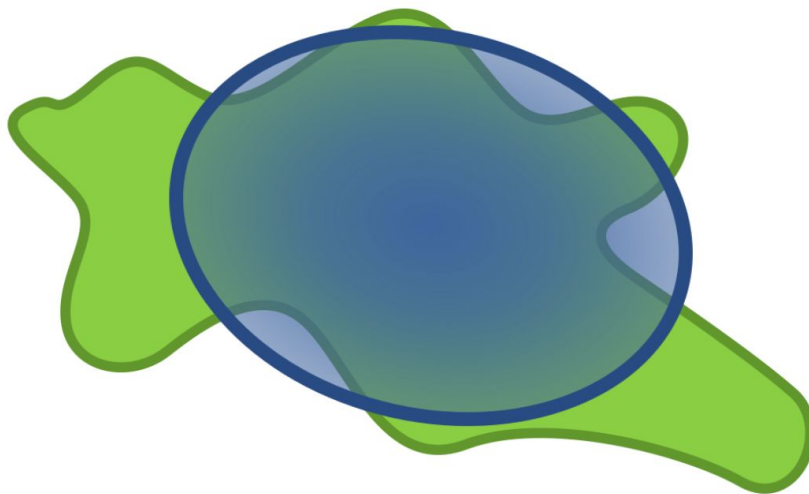
Shift towards end-to-end WSOD.

# Overarching challenge: ill-posedness.

**Weakly-supervised pixel labelling** also requires priors & hints to solve the problem.

## Priors

- Size
- Shape
- Location
- Number of instances
- Contrast (boundaries, saliency)
- Class distribution
- Motion
- Similarity across images
- Similarity with external images

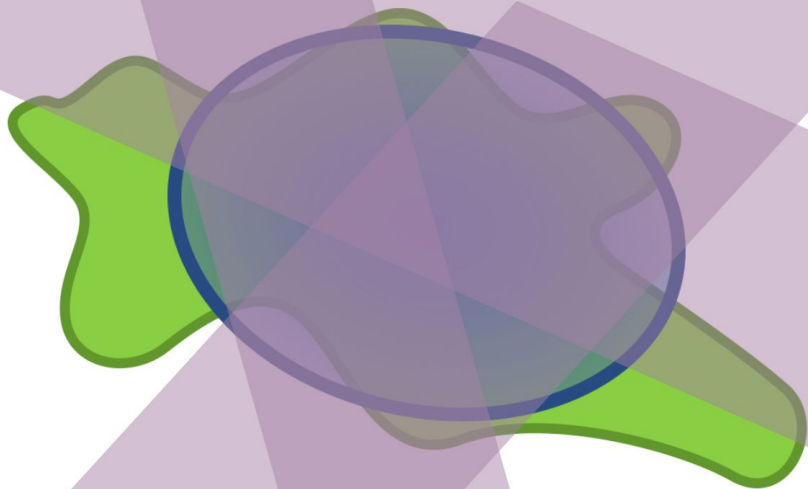


# Overarching challenge: ill-posedness.

**Weakly-supervised pixel labelling** also requires priors & hints to solve the problem.

## Hints

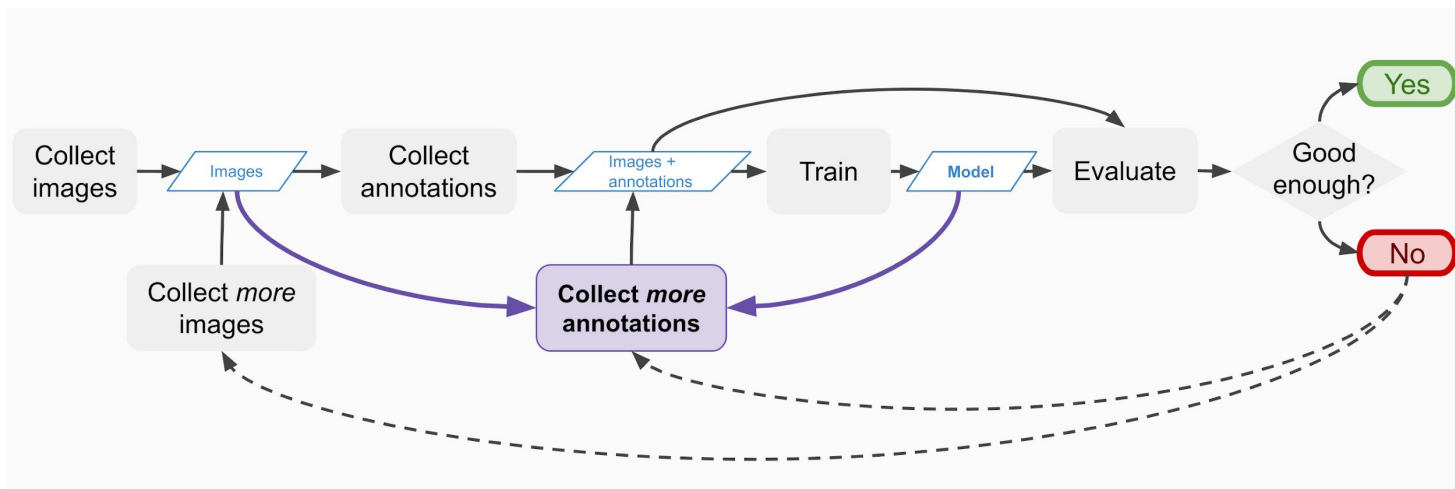
- Image labels
- Image captions
- Video labels
- Transfer across images
- Transfer across classes
- Click inside object
- Size-from-center-click
- Scribbles
- Eye gaze
- Object bounding boxes
- Objects extreme points



# Overarching challenge: ill-posedness.

Weak supervision in practice: **human in the loop**.

Asking for humans' help to resolve the ill-posedness of the problem.



# Overarching challenge: ill-posedness.

**Evaluation problem** in weakly-supervised learning:

- On principle: you are only allowed to use weak supervision.
- But you need extra supervision to really solve the problem.
- Researchers have sought different ways to collect extra supervision.

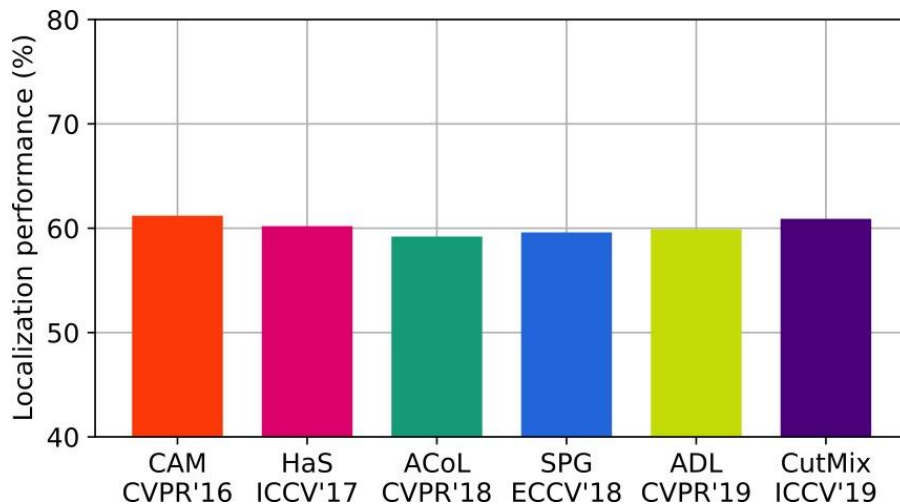
→ Unfair comparisons!



# Overarching challenge: ill-posedness.

Solution: **legalise and control.**

- Fix the type and amount of extra supervision.
- Weakly-supervised object localization methods are re-ranked.





# Overarching challenge: ill-posedness.

Future directions for weakly-supervised learning.

## **Academic benchmarks and evaluation protocols**

- Acknowledge the need for extra priors and hints.
- Include extra information as part of the benchmarks.

## **Methodologies**

- How to combine various levels and types of supervision?
- What are the extra sources of priors and hints?

# Conclusion

- Computer vision requires lots of data; data is expensive.
- Weakly-supervised learning is actively used and required in practice.
- Weakly-supervised learning has many exciting open challenges.
- It is a good time to join the field and make contributions!